

2021-01-01

Comparison of the sensitivity of patient-reported outcomes for detecting the benefit of biologics in severe asthma

Hyland, ME::0000-0003-3879-0469

<http://hdl.handle.net/10026.1/18314>

10.1177/14799731211043530

Chronic Respiratory Disease

SAGE Publications

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

Comparison of the sensitivity of patient reported outcomes for detecting the benefit of biologics in severe asthma

Michael E. Hyland^{1,2}, Joseph W. Lanario², Andrew Menzies-Gow³, Adel H Mansur⁴, James W. Dodd⁵, Stephen J Fowler⁶, Gemma Hayes⁷, Rupert C. Jones², Matthew Masoli⁸

1. Plymouth Marjon University, Plymouth, UK
2. Faculty of Health, University of Plymouth, Plymouth, UK
3. Royal Brompton Hospital, London, UK
4. Heartlands Hospital, University Hospitals Birmingham and University of Birmingham, Birmingham, UK
5. Academic Respiratory Unit, Southmead Hospital, North Bristol Hospital Trust, Bristol, UK
6. School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre and NIHR Manchester Biomedical Research Centre, Manchester University Hospitals NHS Foundation Trust, Manchester, UK
7. University Hospitals Plymouth NHS Trust, Plymouth, UK
8. Royal Devon and Exeter Hospital, University of Exeter, Exeter, UK

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Abstract

Background

The sensitivity of patient reported outcomes (PROs) to detect the effects of treatment change depends on the match between the change in items of the PRO and the change that takes place in a sample of people. The aim of this study is to compare the sensitivity of different PROs in detecting changes following the initiation of biologic treatment in asthma.

Methods

Patients starting a biologic treatment as part of clinical care completed the Asthma Control Questionnaire (ACQ-6), the Severe Asthma Questionnaire (SAQ and SAQ-global scores) and the EQ5D (EQ-5D-5L and EQ5D-VAS) at baseline. They completed the ACQ-6, SAQ, SAQ-global and a retrospective global rating of change (GRoC) scale at weeks 4, 8 and 16, and completed the EQ-5D-5L and EQ5D-VAS at week 16. The SAQ-global and EQ5D-VAS differ but both are single item 100-point questions. Sensitivity was measured by Cohen's D effect size at each of the three time points.

Results

110 patients were recruited. Depending on the time of assessment, effect size varied between 0.45 – 0.64 for the SAQ, between 0.50 – 0.77 for the SAQ-global; between 0.45 – 0.69 for ACQ-6; between 0.91 - 1.22 for GRoC; 0.32 for EQ-5D-5L and 0.49 for EQ5D-VAS.

Conclusion

The sensitivity to change of a questionnaire varies with the time of measurement. The three asthma-specific prospective measures (SAQ, SAQ-global and ACQ-6) have similar sensitivity to change. The single-item EQ5D-VAS was less sensitive than the asthma specific measures and less sensitive than the single-item SAQ-global. The EQ-5D-5L was least sensitive.

Keywords: Asthma; Quality of life; Measurement, SAQ, ACQ-6, EQ5D

Introduction

Patient reported outcomes (PRO) form an increasingly important part of treatment evaluation because they measure the impact of treatment from the patient's perspective. In this paper we compare four different types of PRO in their ability to detect change over time in severe asthma following initiation of biologic therapy: a symptom questionnaire measuring asthma control, two quality of life questionnaires, and a questionnaire of perceived change. A PRO is sensitive to change to the extent that the items of the questionnaire capture the changes that occur in a specific population for a specific treatment [1]. Each of the four types of questionnaire capture different kinds of change and therefore may not be equally sensitive to treatment.

Asthma symptoms are, in part, associated with increased airflow obstruction leading to a lower forced expiratory volume in 1 second (FEV₁). The asthma control questionnaire (ACQ-6) is a validated questionnaire that measures asthma control by the severity of asthma symptoms and shows a strong correlation with lung function [2,3]. The ACQ-6 should therefore provide a sensitive measure of improvements in lung function produced by pharmacological treatment.

Quality of life (QoL) is affected by symptoms but also by other factors such as personality [4], economic status [5] and social support [6]. Small changes in symptoms could create large changes in QoL and the relevant questionnaires, as long as the QoL questionnaire measures aspects of life that change following improved asthma control.

There are two types of quality of life questionnaire, disease specific and generic. As their names suggest, disease specific questionnaires measure only deficits experienced in a particular disease whereas generic questionnaires measure deficits across a range of diseases. The severe asthma questionnaire (SAQ) [7] is designed to measure the QoL deficits that occur in severe asthma and therefore may be more sensitive to change compared to a generic questionnaire such as the EQ5D [8], which is known to poorly represent the deficits of QoL in severe asthma [9] and therefore should be less sensitive to change.

Measures of perceived change, such as global rating of change (GRoC) require people to evaluate how much they feel they have improved. These are seldom used as a primary outcome measures because of concerns about the accuracy of retrospective recall and other possible biases [10]. Retrospective recall of health events is known to be poor, including an underestimation of asthma symptoms [11] and people create implicit narratives of their treatment that can create an over-estimation of the benefit of treatment [10].

The aim of this study is to compare the sensitivity to change as measured by effect size over several time points for four questionnaires: SAQ, ACQ-6, EQ5D and GRoC.

Methods

Design: An open label multi-centre sensitivity to change study, using the SAQ, other questionnaire measures and clinic data.

Participants: Participants were recruited from six UK specialist severe asthma centres (Royal Devon & Exeter, University Hospitals Plymouth NHS Trust, North Bristol NHS Trust, University Hospitals Birmingham, Royal Brompton and Harefield Hospitals, Manchester University NHS Foundation Trust) as part of usual care. All patients had undertaken a full multidisciplinary assessment according to UK severe asthma guidelines [12]. Inclusion criteria: patients with severe asthma (GINA step 4 & 5) commencing a biologic treatment in normal clinical care following NICE guidelines.

Participants were asked to complete questionnaires at baseline and at routine biologic administration clinic visits (0, 4, 8, 12 and 16 weeks).

Questionnaires

Severe Asthma Questionnaire (SAQ): The SAQ is a validated specific scale for severe asthma (10) comprising 16 items measuring difficulties caused by asthma in 16 domains over a 2 week period [7]. Responses are given on a 7-point scale that are averaged to form the total SAQ score and three subscales, SAQ-MyLife, SAQ-MyMind, SAQ-MyBody [13]. In addition, the SAQ provides a SAQ-global score based on a single 100 point Borg-type scale. Higher scores indicate better quality of life. The minimum clinically important difference (MCID) is 0.5 for the SAQ and 11 for the SAQ-global

Asthma Control Questionnaire (ACQ-6): The ACQ-6 has six items, of which four measure symptoms, one measures short acting bronchodilator use and one activity limitations over a 7 day period. Responses are given on a 7-point scale that

averaged to form the ACQ-6 score. Lower scores indicate better asthma control and a score of 0 to 0.75 is classified as well controlled asthma [2]. The MCID is 0.5.

The Euroqol (EQ-5D-5L): The Euroqol is a generic health status questionnaire comprising five dimensions measured on a single day [8]. We used the index values of this scale rather than the average values, as index values are used in resourcing decisions. Index values were calculated using the 2012 value set for England [14].

Global rating of change (GRoC) Questionnaire: Patients were asked to rate the GRoC in terms of improvement by circling a statement 'which best describes how you feel since starting your new treatment for your asthma'. The GRoC used in this study comprises 11 response options, no change, a little better, somewhat better, moderately better, a good deal better and a great deal better, with deterioration represented by the same quantifiers but using the word 'worse' [10].

Procedure:

After providing informed consent, patients completed baseline questionnaires (SAQ and ACQ-6) in the clinic before their first treatment. Patients completed questionnaires (SAQ, ACQ-6, GRoC) during subsequent routine clinic visits as part of usual care at 4 weeks, 8 weeks, 12 weeks and 16 weeks after starting treatment. Patients receiving benralizumab attended at 0, 4, 8 and 16 weeks and so did not attend clinic at 12 weeks due to the transition to an 8 weekly dosing schedule. The EQ-5D-5L and FEV₁ was assessed at baseline and at 16 weeks. Clinical records were used to provide demographic details. Exacerbations requiring OCS, healthcare utilisation and prednisolone dose (if on maintenance OCS) were documented at each clinic visit.

Ethical approval

This study received ethical approvals from the Research Ethics Committee/Health Research Authority (REC reference: 19/WA/0011, IRAS project ID: 250167) and was sponsored by University Hospitals Plymouth NHS Trust.

Statistics

We present the data for baseline, 4, 8 and 16 weeks only as the numbers attending week 12 were reduced as those receiving benralizumab did not attend. The COVID pandemic coincided with the start of data collection so data at 16 weeks was reduced because many of the health staff who were collecting data were reassigned to COVID related duties. For this reason we present two analyses, one intention to treat and one per protocol, and we used t-tests to compare the per protocol patients with the intention to treat after removing the per protocol patients from the intention to treat sample. There is no imputation of missing data for the intention to treat sample: we report all valid data but only valid data. Although the sample size is reduced at later time points the use of the two methods of analysis has the advantage of providing additional calculations of sensitivity to change.

Effect size was Cohen's D calculated from the mean difference in score between baseline and follow up divided by the pooled standard deviation, except in the case of GRoC where Cohen's D was calculated from the mean score (i.e., the difference between the score and zero) divided by the standard deviation of the GRoC score. Cohen's D provides a standardised measure of change between questionnaires and therefore permits comparison of questionnaires that have different scaling properties. Conventional interpretation of Cohen's D is that an effect size of 0.2 is considered small, 0.5 is medium and 0.8 large [15]. We used Spearman's ranked correlations.

Results

One hundred and ten patients were enrolled in the study of whom 62 were female and 48 male. Any patient who completed one or more questionnaires at any time point (0, 4, 8, 16 weeks) was included in the intention to treat sample. Twenty two patients completed all questionnaires at weeks 0, 4, 8, and 16 of whom 12 were female and 10 male and these patients form the per protocol sample. Of the 22 per protocol patients 18 came from one severe asthma centre. Demographic and clinical characteristics of the intention to treat sample are shown in Table 1.

Table 1. Patient characteristics, means (95% confidence intervals)

	Intention to treat sample		Per protocol sample	
	n	Mean (CI)	n	mean
Age, years	106	50 (47 – 53)	22	51 (44 - 59)
Female, (%)	59 (56)			
FEV₁, %predicted	103	66.82 (62.69 – 70.96)	22	65.83 (57.32 - 74.34)
Caucasian, (%)	91 (86)		20 (91)	
BMI, kg/m²	104	29.49 (28.06 – 30.92)	22	28.20 (25.85 – 30.54)
Maintenance prednisolone dose, mg/day	60	12.43 (9.64 – 15.21)	11	10.70 (5.90 – 5.43)
Exacerbations in the last 12 months requiring OCS	106	5.43 (4.54 – 6.31)	22	6.50 (4.23 – 8.77)
Emergency Department visits	106	2.08 (1.09 – 3.06)	22	2.27 (0.58 – 5.13)
Hospital admissions	106	1.25 (0.48 – 2.01)	22	0.64 (0.16 -1.26)
Cumulative prednisolone dose, mg/yr	105	4062 (3238 – 4885)	22	3766 (2472 - 5060)
Receiving biologic treatment, n (%)	106 (96)			
Omalizumab (%)	16 (15)		4 (18)	
Mepolizumab (%)	26 (25)		10 (45)	
Benralizumab (%)	62 (59)		8 (36)	
Reslizumab (%)	2 (2)		0	

Table 2 shows the mean and standard deviations for the SAQ, SAQ-global and ACQ-6 at baseline, 4, 8 and 16 weeks, for the EQ-5D-5L at baseline and 16 weeks, and for the GROC at 4, 8, and 16 weeks. The data are shown separately for intention

to treat and per protocol and there was no significant difference between the intention to treat and per protocol group at any time point.

Table 2. **Means** and *standard deviations* for questionnaires for the intention to treat (ITT) group (n) and the per protocol (PP) group (n=22).

	Baseline		4 weeks		8 weeks		16 weeks	
	ITT	PP	ITT	PP	ITT	PP	ITT	PP
SAQ	3.72 (106) 1.46	3.58 1.21	4.38 (102) 1.54	4.40 1.45	4.90 (73) 1.51	4.54 1.52	4.57 (48) 1.43	4.72 1.35
SAQ-Global	46.13 (106) 23.67	47.72 19.13	57.19 (105) 22.72	59.31 19.90	60.40 (73) 24.70	67.45 18.35	63.15 (47) 23.79	63.64 19.22
ACQ-6	3.13 (90) 1.40	3.07 1.08	2.53 (85) 1.50	2.30 1.24	2.35 (59) 1.46	2.02 1.12	2.23 (41) 1.53	2.17 1.28
EQ5D index	0.69 (103) 0.23	0.71 0.19					0.76 (41) 0.21	0.80 0.18
EQ5D-VAS	54.76 (101) 22.21	56.82 19.43					65.20 (36) 23.35	70.65* 18.43
GRoC			1.83 (103) 2.01	2.09 2.07	2.25 (72) 2.40	3.00 1.66	2.84 (45) 2.36	3.18 1.68

Table 3 shows the effect sizes at the three time points after baseline for all questionnaires as a function of the intention to treat and per protocol. Of the six comparisons made between the asthma specific questionnaires, the SAQ-global is slightly better than the other two for four comparisons and the SAQ slightly better than the other two for two comparisons. Overall, though variable the relative effect sizes are similar over time and between groups. However, the effect sizes of the EQ-5D-5L and EQ5D-VAS are lower than the disease specific scale for all comparisons.

Table 3. Effect sizes (Cohen's D) for change in questionnaire scores at three follow-up time points for intention to treat (ITT) and per protocol (PP) samples.

	4 weeks		8 weeks		16 weeks	
	ITT	PP	ITT	PP	ITT	PP
SAQ	0.44	0.61	0.38	0.70	0.60	0.89
SAQ-Global	0.48	0.59	0.60	1.06	0.72	0.83
ACQ-6	0.41	0.58	0.54	0.95	0.61	0.76
EQ5D					0.32	0.49
EQ5D-VAS					0.49	0.73
GRoC	0.91	1.01	0.94	1.81	1.20	1.89

Table 4 shows the baseline and follow up scores for the subscales of the SAQ as a function of the intention to treat and per protocol, and shows similar performance across all comparisons.

Table 4. **Means** and *standard deviations* for SAQ subscale scores for the intention to treat (ITT) group (n) and the per protocol (PP) group (n=22).

	Baseline		4 weeks		8 weeks		16 weeks	
	ITT	PP	ITT	PP	ITT	PP	ITT	PP
SAQ My Life	3.80 (104) 1.60	3.75 1.37	4.45 (100) 1.70	4.48 1.59	4.37 (71) 1.70	4.70 1.54	4.68 (47) 1.56	4.82 1.33
SAQ My Mind	3.87 (104) 1.74	3.63 1.45	4.49 (100) 1.72	4.50 1.62	4.29 (71) 1.73	4.42 1.78	4.69 (47) 1.58	4.82 1.65
SAQ My Body	3.42 (104) 1.48	3.09 1.27	4.06 (100) 1.58	4.06 1.53	4.01 (71) 1.43	4.27 1.52	4.28 (47) 1.50	4.38 1.50

Table 5 shows the effect sizes at the three time points after baseline for the subscales of the SAQ as a function of the intention to treat and per protocol, and shows similar performance across all comparisons.

Table 5. Effect sizes (Cohen's D) for change in SAQ subscale scores at three follow-up time points.

	4 weeks		8 weeks		16 weeks	
	ITT	PP	ITT	PP	ITT	PP
SAQ My Life	0.39	0.49	0.35	0.65	0.56	0.79
SAQ My Mind	0.36	0.57	0.24	0.49	0.49	0.77
SAQ My Body	0.42	0.69	0.41	0.84	0.58	0.97

Table 6 shows the correlations between the questionnaires and objective measures at baseline

Table 6. Correlation coefficients (rho) between questionnaire scores at baseline and objective measures.

	FEV1 percent predicted	Exacerbations in the last 12 months requiring OCS
SAQ	0.25*	-0.34**
SAQ-global	0.27**	-0.27**
SAQ-My Life	0.30**	-0.35**
SAQ-My Mind	0.14	-0.27**
SAQ-My Body	0.13	-0.29**
ACQ-6	-0.31**	0.32*
EQ-5D-5L	0.28**	-0.28**
EQ5D-VAS	0.34**	-0.30

Discussion

The sensitivity of a questionnaire to change depends on the items of the questionnaire, on the treatment and on the population studied [1]. In this study sensitivity was assessed at three time points. The dropout rate for this study was high due to the pandemic, and as dropout can be non-random we provided two analysis, intention to treat and per protocol. Although there was no significant difference there is a trend for those in the per protocol to have higher effect sizes compared to those in the intention to treat sample. For both the intention to treat and

per protocol groups, the MCID for the SAQ, SAQ-global and ACQ-6 was achieved as early as week 4.

As we performed a per protocol analysis in addition to an intention to treat analysis, it is possible to compare the relative effect sizes of the questionnaires in different samples. Our results reveal that the relative effect size between questionnaires varies as a function of the population and as a function of time. Sensitivity to change is not an absolute property of a scale but something that varies over time and with the population. Previous comparisons of the sensitivity of quality of life questionnaires have provided comparisons at only one time point and with only one group [16,17]. The variability of effect size as a function of group and time should be recognised when comparing between questionnaires.

Our results provide six comparisons of effect size. Some questionnaires are consistently better or worse than others. Some questionnaires are sometimes better or worse depending on which comparison is made.

The results show that that the relative sensitivity, as measured by effect size, of the different questionnaires changes to some extent between the six different comparisons, showing that comparisons of sensitivity drawn from only one group and at one time may not be generalisable. There is no evidence of important differences in sensitivity between the SAQ, SAQ-global and ACQ-6. The SAQ-global is slightly more sensitive on four of the six comparisons but, given the variability in relative effect sizes observed between these questionnaires, there is no clear evidence of difference between them.

The effect size of the EQ-5D-5L is consistently less than all the other questionnaires in all comparisons. The effect size of the EQ-5D-5L is about half that of the SAQ, SAQ-global and ACQ-6 and it is therefore safe to conclude that the EQ-5D-5L is the least sensitive of the questionnaires studied. The comparative insensitivity of the EQ-5D-5L is to be expected from a scale that only partially captures the QoL deficits of severe asthma on a single day [8] and findings are also consistent with evidence that this scale is comparatively insensitive to rehabilitation [18].

The sensitivity of the EQ5D-VAS was slightly better than the EQ-5D-5L, but not as sensitive as other scales and, in particular, less sensitive than the SAQ-global.

Although the EQ5D-VAS and SAQ-global are single item scales varying between 0 and 100, they differ in two respects. The EQ5D-VAS asks patients to rate health, the SAQ-global asks patients to rate quality of life specifically in relation to asthma. The EQ-5D-VAS is a scale with only the end points specified. The SAQ-global is a Borg scale with additional quantifiers placed at empirically derived points along the scale. Borg scales have been shown to be more reliable than visual analogue scales [19]. Our data shows that they may also be more sensitive.

The SAQ has three subscales made up from groupings of the 16 items that make up the SAQ. There was no consistent difference in sensitivity for these three subscales.

The My Body subscale was the most sensitive of the three at all three time points for the per protocol group but not for the intention to treat subscale, where it was most sensitive only at week 8. These findings illustrate how small differences in sensitivity can arise from differences in population and time point, again showing that sensitivity is not an absolute property of a questionnaire, but relative to the population and time of assessment.

Although all three subscales of the SAQ are sensitive to the effects of treatment, only the My Life subscale correlates with baseline FEV₁%. A larger data set has also shown that the correlation between FEV₁% and the My Life subscale is larger than My Life and the My Body subscales [13]. As patients' judgements of quality of life are determined by numerous factors (including dispositional mood, lung function, effects of treatment) these data suggest that the My Life and My Body subscales are more affected by factors other than lung function. Examination of other correlations with other PROs demonstrated no relationship between sensitivity to change and FEV₁%. The conclusion from these data is that the use of QoL as an outcome variable provides different information from change in lung function, and that both PROs and objective measures should be used in clinical trials as they provide different kinds of information about the effects of treatment.

Our study measured change in an open label study rather than comparative change between placebo and active treatment. We found that a global rating of change scale (GRoC) had the highest sensitivity to change, but it does not follow that it is most sensitive in comparing placebo with active treatment. The reason is that retrospective measures are known to be affected by biases, including that of implicit theory [20]. When patients receive a new treatment, they form a narrative that the treatment is likely to be effective, and their response, for both placebo and treatment is therefore affected by this narrative [21,22].

The high sensitivity of the SAQ-global but not the EQ5D-VAS shows that single item scales can sometimes be highly sensitive. The SAQ-global is derived from an earlier scale, the Global Quality of Life Scale (GQoL) [18]. Multi-item scales require the

patient to consider the components of quality of life. These components are then aggregated most commonly with no weighting, but sometimes as in the case of the EQ5D with weightings from people who are not ill. As a result, the components of the multi-item quality of life scale are not aggregated using the patient's own individual utilities or weights. By contrast, the SAQ-global requires patients do this aggregation themselves, something that will take place fast, automatically, and with implicit rather than conscious consideration of events in a person's life [23]. Our findings indicate that this fast, implicit process can be no worse than that provided by longer questionnaires in detecting change.

Conclusion

The sensitivity of the ACQ-6, SAQ and SAQ-global to the effect of starting biologic treatment is similar. The EQ5D-VAS was less sensitive and the EQ-5D-5L was least sensitive to treatment. The comparative sensitivity of PROs varies slightly as a function of population and time of assessment, so comparisons with other data sets should be treated with caution.

Our study shows that in clinical practice a rapid increase in quality of life can be detected by the SAQ and SAQ-global. The easy use of the SAQ-global makes it a suitable tool if time is limited. The EQ-5D should not be used in clinical practice to evaluate change in quality of life in severe asthma as it is comparatively insensitive to change, and the use of this questionnaire in economic decision making may lead to an underestimation of the benefit of biologic treatment.

Acknowledgements

This research was made possible by a non-promotional grant provided by GSK. JL is supported by the National Institute for Health Research (NIHR) Applied Research Collaboration (ARC) South West Peninsula. SF is supported by NIHR Manchester Biomedical Research Centre and the study supported by the NIHR Manchester Clinical Research Facility. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

References

1. Hyland ME. Selection of items and avoidance of bias in quality of life scales. *Pharmacoeconomics*. 1992 Mar;1(3):182-90.
2. Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. *European Respiratory Journal*. 1999;14:902-7.
3. Thamrin C, Nydegger R, Stern G, Chanez P, Wenzel SE, Watt RA, FitzPatrick S, Taylor DR, Frey U: Associations between fluctuations in lung function and asthma control in two populations with differing asthma severity. *Thorax* 2011, 66:1036-1042.
4. Huang IC, Lee JL, Ketheeswaran P, Jones CM, Revicki DA, Wu AW. Does personality affect health-related quality of life? A systematic review. *PloS one*. 2017 Mar 29;12(3):e0173806..
5. Niedzwiedz CL, Katikireddi SV, Pell JP, Mitchell R. Life course socio-economic position and quality of life in adulthood: a systematic review of life course models. *BMC public health*. 2012 Dec;12(1):1-1.

6. Helgeson VS. Social support and quality of life. *Quality of life research*. 2003 Jan;12(1):25-31.
7. Hyland ME, Jones RC, Lanario JW, Masoli M. The construction and validation of the Severe Asthma Questionnaire. *European Respiratory Journal*. 2018 Jul 1;52: 1800618
8. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, Swinburn P, Busschbach J. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Quality of Life Research*. 2013 Sep 1;22(7):1717-27.
9. Whalley D, Globe G, Crawford R, Doward L, Tafesse E, Brazier J, Price D. Is the EQ-5D fit for purpose in asthma? Acceptability and content validity from the patient perspective. *Health and quality of life outcomes*. 2018 Dec;16(1):1-4.
10. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *Journal of Manual & Manipulative Therapy*. 2009;17(3):163-70.
11. Hyland ME, Kenyon CA, Allen R, Howarth P. Diary keeping in asthma: comparison of written and electronic methods. *British Medical Journal*. 1993 Feb 20;306(6876):487-9.
12. England N. Specialised Respiratory Services (adult)–Severe Asthma. 2018.
13. Lanario JW, Hyland ME, Menzies-Gow A, Mansur AH, Dodd JW, Fowler SJ, Jones RC, Masoli M. Validation of subscales of the Severe Asthma Questionnaire (SAQ) using exploratory factor analysis (EFA). *Health and quality of life outcomes*. 2020 Dec;18(1):1-0.

14. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: An EQ-5 D-5 L value set for England. *Health economics*. 2018 Jan;27(1):7-22.
15. Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
16. Rutten-van Molken MP, Custers F, Van Doorslaer EK, Jansen CC, Heurman L, Maesen FP, Smeets JJ, Bommer AM, Raaijmakers JA. Comparison of performance of four instruments in evaluating the effects of salmeterol on asthma quality of life. *European Respiratory Journal*. 1995 Jun 1;8(6):888-98.
17. Sanjuás C, Alonso J, Prieto L, Ferrer M, Broquetas JM, Antó JM. Health-related quality of life in asthma: a comparison between the St George's Respiratory Questionnaire and the Asthma Quality of Life Questionnaire. *Quality of Life Research*. 2002 Dec;11(8):729-38.
18. Szentes BL, Schultz K, Nowak D, Schuler M, Schwarzkopf L. How does the EQ-5D-5L perform in asthma patients compared with an asthma-specific quality of life questionnaire?. *BMC pulmonary medicine*. 2020 Dec;20(1):1-1.
19. Hyland ME, Sodergren SC. Development of a new type of global quality of life scale, and comparison of performance and preference for 12 global scales. *Quality of Life Research*. 1996 Oct;5(5):469-80.
20. Norman, G. (2003). Hi! How are you? Response shift, implicit theories and differing epistemologies. *Quality of Life Research*, 12, 239–249.
21. Robinson J, Khan N, Fusco L, Malpass A, Lewis G, Dowrick C. Why are there discrepancies between depressed patients' Global Rating of Change and scores on the Patient Health Questionnaire depression module? A qualitative study of primary care in England. *BMJ open*. 2017 Apr 1;7(4).

22. Schmitt J, Di Fabio RP. The validity of prospective and retrospective global change criterion measures. Archives of physical medicine and rehabilitation. 2005 Dec 1;86(12):2270-6.
23. Kahneman D. Thinking, fast and slow. Macmillan; 2011 Oct 25.